Final Abstract

Anaerobic digestion is a critical renewable-energy pathway for converting organic waste into methane-rich biogas, yet reliable methane-yield prediction remains difficult because the process is nonlinear, time-dependent, biologically sensitive, and strongly affected by interacting variables such as temperature, pH, organic loading rate, hydraulic retention time, total solids, volatile solids, C/N ratio, moisture, and prior reactor state. Mechanistic models such as ADM1 provide scientific interpretability but are complex to calibrate for real-time decision support, while purely data-driven models can produce predictions that are accurate statistically but weak from a process-feasibility perspective.

This dissertation presents the Green Fuel Intelligence System (GFIS), a physics-guided hybrid AI digital twin for methane-yield prediction, stability monitoring, simulation, and operational optimization in anaerobic digestion. GFIS integrates an end-to-end data pipeline, engineered process features, an XGBoost tabular predictor, a PyTorch LSTM time-series model, an ensemble prediction layer, a volatile-solids-based methane feasibility constraint, and a VFA/ALK soft sensor for classifying digester stability as Stable, Warning, or Critical. The system is implemented as a working software prototype with FastAPI services, SQLite logging, Streamlit interaction, and browser-based digital-twin demonstrations for both model-backed control and industrial 48-hour simulation.

The current prototype is validated on a synthetic and literature-inspired anaerobic-digestion dataset designed to support controlled experimentation before real plant integration. On the evaluation set, the XGBoost model achieved an R2 score of 0.7556, RMSE of 6.7933, and MAE of 5.5517, while the physics-guided layer reported zero infeasible methane-yield violations. Additional 48-hour stress experiments demonstrate GFIS behavior under nominal operation, organic overload, pH shock, low-temperature disturbance, reduced HRT, high-solids feedstock, and corrective recovery. Organic overload correctly moved the soft sensor into a Warning condition, while all tested scenarios remained physically bounded.

The major contribution of this work is not only methane prediction, but a complete research-grade decision-support framework that combines machine learning accuracy, physics consistency, stability awareness, explainability, simulation, optimization, and deployable software architecture. GFIS establishes a practical foundation for future calibration with plant data, IoT/SCADA integration, PostgreSQL-backed deployment, and advanced cyber-physical simulation using ROS2/Gazebo. The dissertation contributes an AI-driven pathway for safer, more interpretable, and more operationally useful biogas process intelligence.

Keywords: Anaerobic digestion, methane yield prediction, physics-guided machine learning, digital twin, XGBoost, LSTM, VFA/ALK soft sensor, biogas optimization, process stability, sustainable energy.